

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/587,956
Source: IFWP
Date Processed by STIC: 08/14/2006

ENTERED



IFWP

RAW SEQUENCE LISTING

DATE: 08/14/2006

PATENT APPLICATION: US/10/587,956

TIME: 13:18:35

Input Set : A:\Sequence Listing.TXT

Output Set: N:\CRF4\08142006\J587956.raw

```

1 <110> APPLICANT: Korea Research Institute of Bioscience and Biotechnology
3 <120> TITLE OF INVENTION: A novel Hansenula polymorpha gene coding for alpha 1,6
4     mannosyltransferase and process for the production of recombinant
5     glycoproteins with Hansenula polymorpha mutant strain deficient
6     in the same gene
C--> 8 <140> CURRENT APPLICATION NUMBER: US/10/587,956
C--> 8 <141> CURRENT FILING DATE: 2006-07-31
W--> 0 <130> FILE REFERENCE:
      8 <160> NUMBER OF SEQ ID NOS: 14
10 <170> SOFTWARE: KopatentIn 1.71
12 <210> SEQ ID NO: 1
13 <211> LENGTH: 1351
14 <212> TYPE: DNA
15 <213> ORGANISM: Hansenula polymorpha
17 <220> FEATURE:
18 <221> NAME/KEY: CDS
19 <222> LOCATION: (10)..(1293)
21 <400> SEQUENCE: 1
22 cggtgaaga  atg gtg tat ttt tta aat ttc atg tca ata acc aat gtc ccg      51
23           Met Val Tyr Phe Leu Asn Phe Met Ser Ile Thr Asn Val Pro
24           1               5               10
26 gtg ctg aag cgc gcg cga ctc tac atg gcg acg aat cgc cgg ctg gtg      99
27 Val Leu Lys Arg Ala Arg Leu Tyr Met Ala Thr Asn Arg Arg Leu Val
28 15           20           25           30
30 gtt gtt ctt gtg gtg ctg ctg tac tgg gtg gtc cag aac gtt tgg acg      147
31 Val Val Leu Val Val Leu Leu Tyr Trp Val Val Gln Asn Val Trp Thr
32           35           40           45
34 tgg agc cct ggg acg cgc gat ttg gcc caa gtg gac gcg aag atc gag      195
35 Trp Ser Pro Gly Thr Arg Asp Leu Ala Gln Val Asp Ala Lys Ile Glu
36           50           55           60
38 gcc gag cta aac tcg aat cta cat act ttt gga gcg cat ttg cgc cac      243
39 Ala Glu Leu Asn Ser Asn Leu His Thr Phe Gly Ala His Leu Arg His
40           65           70           75
42 tta aac cgg ctt ccg gca gag tcg gcc acc ctg cgt gaa aaa ctc acc      291
43 Leu Asn Arg Leu Pro Ala Glu Ser Ala Thr Leu Arg Glu Lys Leu Thr
44           80           85           90
46 ttc tat ttc cca tat tat cct gaa aag ccc gtg ccg aac cag atc tgg      339
47 Phe Tyr Phe Pro Tyr Tyr Pro Glu Lys Pro Val Pro Asn Gln Ile Trp
48 95           100          105          110
50 cag aca tgg aag gtc gat ctc gaa gac gac aac ttc ccc aag cag tac      387
51 Gln Thr Trp Lys Val Asp Leu Glu Asp Asp Asn Phe Pro Lys Gln Tyr
52           115          120          125
54 aga cgg ttt cag aag acg tgg gtc gag aaa aat cca gac tac gtg tac      435

```

RAW SEQUENCE LISTING

DATE: 08/14/2006

PATENT APPLICATION: US/10/587,956

TIME: 13:18:35

Input Set : A:\Sequence Listing.TXT

Output Set: N:\CRF4\08142006\J587956.raw

55	Arg	Arg	Phe	Gln	Lys	Thr	Trp	Val	Glu	Lys	Asn	Pro	Asp	Tyr	Val	Tyr	
56				130					135					140			
58	cac	ctg	att	ccg	gac	tct	gtg	att	gag	gac	ttt	gtg	gcg	agt	ttg	tac	483
59	His	Leu	Ile	Pro	Asp	Ser	Val	Ile	Glu	Asp	Phe	Val	Ala	Ser	Leu	Tyr	
60				145					150					155			
62	gcg	aac	gtg	ccg	gag	gtg	gtc	aga	gcg	tac	cag	ctg	ctt	ccg	aaa	aat	531
63	Ala	Asn	Val	Pro	Glu	Val	Val	Arg	Ala	Tyr	Gln	Leu	Leu	Pro	Lys	Asn	
64				160					165					170			
66	atc	atg	aag	gcg	gat	ttt	ttc	cgg	tat	ttg	gtg	atc	tac	gcg	cgc	gga	579
67	Ile	Met	Lys	Ala	Asp	Phe	Phe	Arg	Tyr	Leu	Val	Ile	Tyr	Ala	Arg	Gly	
68	175								180					185		190	
70	ggc	acc	tac	tca	gac	atg	gac	acg	gtg	tgt	tta	aag	ccg	atc	aag	gac	627
71	Gly	Thr	Tyr	Ser	Asp	Met	Asp	Thr	Val	Cys	Leu	Lys	Pro	Ile	Lys	Asp	
72									195					200		205	
74	tgg	gcc	acg	ttt	gat	cgc	gac	ctg	atc	cac	gct	gcc	gac	aat	aag	gcc	675
75	Trp	Ala	Thr	Phe	Asp	Arg	Asp	Leu	Ile	His	Ala	Ala	Asp	Asn	Lys	Ala	
76				210					215					220			
78	gat	ctc	tcc	cag	ata	gat	cca	gaa	gca	aga	acc	acg	cct	gtg	ggg	ctg	723
79	Asp	Leu	Ser	Gln	Ile	Asp	Pro	Glu	Ala	Arg	Thr	Thr	Pro	Val	Gly	Leu	
80				225					230					235			
82	gtg	att	ggc	att	gag	gcc	gac	ccg	gac	agg	ccc	gac	tgg	cac	gag	tgg	771
83	Val	Ile	Gly	Ile	Glu	Ala	Asp	Pro	Asp	Arg	Pro	Asp	Trp	His	Glu	Trp	
84				240					245					250			
86	ttc	tcg	cgc	aga	ctg	cag	ttc	tgc	cag	tgg	acg	atc	cag	gcg	aag	ccg	819
87	Phe	Ser	Arg	Arg	Leu	Gln	Phe	Cys	Gln	Trp	Thr	Ile	Gln	Ala	Lys	Pro	
88	255								260					265		270	
90	gga	cac	ccg	ctg	ctg	cgc	gag	ctg	atc	atc	cgg	atc	gtg	gag	gag	acg	867
91	Gly	His	Pro	Leu	Leu	Arg	Glu	Leu	Ile	Ile	Arg	Ile	Val	Glu	Glu	Thr	
92									275					280		285	
94	ttc	cgc	aaa	cag	cac	atg	ggc	gtt	ttg	aaa	aga	gtg	gaa	ggc	aag	gac	915
95	Phe	Arg	Lys	Gln	His	Met	Gly	Val	Leu	Lys	Arg	Val	Glu	Gly	Lys	Asp	
96				290					295					300			
98	tcg	ggc	gca	gat	atc	atg	cag	tgg	aca	gga	ccg	ggg	ata	ttt	aca	gac	963
99	Ser	Gly	Ala	Asp	Ile	Met	Gln	Trp	Thr	Gly	Pro	Gly	Ile	Phe	Thr	Asp	
100				305					310					315			
102	act	ctg	ttt	gat	tat	ctg	aac	aat	gtg	gcg	agc	gac	ggc	aag	ttg	ggc	1011
103	Thr	Leu	Phe	Asp	Tyr	Leu	Asn	Asn	Val	Ala	Ser	Asp	Gly	Lys	Leu	Gly	
104				320					325					330			
106	gac	ggg	tac	ggc	gtg	ggg	tcg	ttg	tat	tgg	cgc	aag	cac	ggc	aaa	tat	1059
107	Asp	Gly	Tyr	Gly	Val	Gly	Ser	Leu	Tyr	Trp	Arg	Lys	His	Gly	Lys	Tyr	
108	335								340					345		350	
110	aag	ctg	aaa	aag	aca	gaa	att	aac	aag	aat	aac	gag	cca	ttg	cat	tct	1107
111	Lys	Leu	Lys	Lys	Thr	Glu	Ile	Asn	Lys	Asn	Asn	Glu	Pro	Leu	His	Ser	
112									355					360		365	
114	gag	gac	cag	ctt	atc	aac	tgg	agg	tcg	ctg	acc	aac	atg	gac	aag	cca	1155
115	Glu	Asp	Gln	Leu	Ile	Asn	Trp	Arg	Ser	Leu	Thr	Asn	Met	Asp	Lys	Pro	
116				370					375					380			
118	aag	atc	atg	ggg	gac	gta	atg	gtg	tta	cca	atc	acg	agc	ttt	agt	ccg	1203
119	Lys	Ile	Met	Gly	Asp	Val	Met	Val	Leu	Pro	Ile	Thr	Ser	Phe	Ser	Pro	

RAW SEQUENCE LISTING

DATE: 08/14/2006

PATENT APPLICATION: US/10/587,956

TIME: 13:18:35

Input Set : A:\Sequence Listing.TXT

Output Set: N:\CRF4\08142006\J587956.raw

```

120          385          390          395
122 aac gtg ggg cac atg ggc tca aag agc agc tca gat agg ctg gca ttt      1251
123 Asn Val Gly His Met Gly Ser Lys Ser Ser Ser Asp Arg Leu Ala Phe
124          400          405          410
126 gtg gag cat tta ttt tct ggc agc tgg aag cca aaa aac aaa   taggaaa      1300
127 Val Glu His Leu Phe Ser Gly Ser Trp Lys Pro Lys Asn Lys
128 415          420          425
130 aataaataat tagctgcatt ttagataatt ctcatgagca ggcacagaac g      1351
133 <210> SEQ ID NO: 2
134 <211> LENGTH: 428
135 <212> TYPE: PRT
136 <213> ORGANISM: Hansenula polymorpha
138 <400> SEQUENCE: 2
139 Met Val Tyr Phe Leu Asn Phe Met Ser Ile Thr Asn Val Pro Val Leu
140 1          5          10          15
142 Lys Arg Ala Arg Leu Tyr Met Ala Thr Asn Arg Arg Leu Val Val Val
143          20          25          30
145 Leu Val Val Leu Leu Tyr Trp Val Val Gln Asn Val Trp Thr Trp Ser
146          35          40          45
148 Pro Gly Thr Arg Asp Leu Ala Gln Val Asp Ala Lys Ile Glu Ala Glu
149          50          55          60
151 Leu Asn Ser Asn Leu His Thr Phe Gly Ala His Leu Arg His Leu Asn
152 65          70          75          80
154 Arg Leu Pro Ala Glu Ser Ala Thr Leu Arg Glu Lys Leu Thr Phe Tyr
155          85          90          95
157 Phe Pro Tyr Tyr Pro Glu Lys Pro Val Pro Asn Gln Ile Trp Gln Thr
158          100          105          110
160 Trp Lys Val Asp Leu Glu Asp Asp Asn Phe Pro Lys Gln Tyr Arg Arg
161          115          120          125
163 Phe Gln Lys Thr Trp Val Glu Lys Asn Pro Asp Tyr Val Tyr His Leu
164          130          135          140
166 Ile Pro Asp Ser Val Ile Glu Asp Phe Val Ala Ser Leu Tyr Ala Asn
167 145          150          155          160
169 Val Pro Glu Val Val Arg Ala Tyr Gln Leu Leu Pro Lys Asn Ile Met
170          165          170          175
172 Lys Ala Asp Phe Phe Arg Tyr Leu Val Ile Tyr Ala Arg Gly Gly Thr
173          180          185          190
175 Tyr Ser Asp Met Asp Thr Val Cys Leu Lys Pro Ile Lys Asp Trp Ala
176          195          200          205
178 Thr Phe Asp Arg Asp Leu Ile His Ala Ala Asp Asn Lys Ala Asp Leu
179          210          215          220
181 Ser Gln Ile Asp Pro Glu Ala Arg Thr Thr Pro Val Gly Leu Val Ile
182 225          230          235          240
184 Gly Ile Glu Ala Asp Pro Asp Arg Pro Asp Trp His Glu Trp Phe Ser
185          245          250          255
187 Arg Arg Leu Gln Phe Cys Gln Trp Thr Ile Gln Ala Lys Pro Gly His
188          260          265          270
190 Pro Leu Leu Arg Glu Leu Ile Ile Arg Ile Val Glu Glu Thr Phe Arg
191          275          280          285

```

RAW SEQUENCE LISTING

DATE: 08/14/2006

PATENT APPLICATION: US/10/587,956

TIME: 13:18:35

Input Set : A:\Sequence Listing.TXT

Output Set: N:\CRF4\08142006\J587956.raw

```

193 Lys Gln His Met Gly Val Leu Lys Arg Val Glu Gly Lys Asp Ser Gly
194      290      295      300
196 Ala Asp Ile Met Gln Trp Thr Gly Pro Gly Ile Phe Thr Asp Thr Leu
197 305      310      315      320
199 Phe Asp Tyr Leu Asn Asn Val Ala Ser Asp Gly Lys Leu Gly Asp Gly
200      325      330      335
202 Tyr Gly Val Gly Ser Leu Tyr Trp Arg Lys His Gly Lys Tyr Lys Leu
203      340      345      350
205 Lys Lys Thr Glu Ile Asn Lys Asn Asn Glu Pro Leu His Ser Glu Asp
206      355      360      365
208 Gln Leu Ile Asn Trp Arg Ser Leu Thr Asn Met Asp Lys Pro Lys Ile
209      370      375      380
211 Met Gly Asp Val Met Val Leu Pro Ile Thr Ser Phe Ser Pro Asn Val
212 385      390      395      400
214 Gly His Met Gly Ser Lys Ser Ser Ser Asp Arg Leu Ala Phe Val Glu
215      405      410      415
217 His Leu Phe Ser Gly Ser Trp Lys Pro Lys Asn Lys
218      420      425
221 <210> SEQ ID NO: 3
222 <211> LENGTH: 36
223 <212> TYPE: DNA
224 <213> ORGANISM: Artificial Sequence
226 <220> FEATURE:
227 <223> OTHER INFORMATION: primer 168Not-N
230 <400> SEQUENCE: 3
231 aaggaaaaaa gcggccgccg gtgaagaatg gtgtat
234 <210> SEQ ID NO: 4
235 <211> LENGTH: 39
236 <212> TYPE: DNA
237 <213> ORGANISM: Artificial Sequence
239 <220> FEATURE:
240 <223> OTHER INFORMATION: primer 168Not-C
243 <400> SEQUENCE: 4
244 ttttcctttt gcggccgccg ttctgtgcct gctcatgat
247 <210> SEQ ID NO: 5
248 <211> LENGTH: 20
249 <212> TYPE: DNA
250 <213> ORGANISM: Artificial Sequence
252 <220> FEATURE:
253 <223> OTHER INFORMATION: primer UNfor
256 <400> SEQUENCE: 5
257 ggatccccgg gtaccgagct
260 <210> SEQ ID NO: 6
261 <211> LENGTH: 20
262 <212> TYPE: DNA
263 <213> ORGANISM: Artificial Sequence
265 <220> FEATURE:
266 <223> OTHER INFORMATION: primer UNrev
269 <400> SEQUENCE: 6

```

36

39

20

RAW SEQUENCE LISTING

DATE: 08/14/2006

PATENT APPLICATION: US/10/587,956

TIME: 13:18:35

Input Set : A:\Sequence Listing.TXT

Output Set: N:\CRF4\08142006\J587956.raw

```

270 caccggtagc taatgatccc                                20
273 <210> SEQ ID NO: 7
274 <211> LENGTH: 20
275 <212> TYPE: DNA
276 <213> ORGANISM: Artificial Sequence
278 <220> FEATURE:
279 <223> OTHER INFORMATION: primer UCfor
282 <400> SEQUENCE: 7
283 cgaacatcca agtgggccga                                20
286 <210> SEQ ID NO: 8
287 <211> LENGTH: 20
288 <212> TYPE: DNA
289 <213> ORGANISM: Artificial Sequence
291 <220> FEATURE:
292 <223> OTHER INFORMATION: primer UCrev
295 <400> SEQUENCE: 8
296 ctggcgaaaag ggggatgtgc                                20
299 <210> SEQ ID NO: 9
300 <211> LENGTH: 20
301 <212> TYPE: DNA
302 <213> ORGANISM: Artificial Sequence
304 <220> FEATURE:
305 <223> OTHER INFORMATION: primer 168Nfor
308 <400> SEQUENCE: 9
309 ggcggatatg gggcttcgcc                                20
312 <210> SEQ ID NO: 10
313 <211> LENGTH: 40
314 <212> TYPE: DNA
315 <213> ORGANISM: Artificial Sequence
317 <220> FEATURE:
318 <223> OTHER INFORMATION: primer 168Nrev
321 <400> SEQUENCE: 10
322 agctcgggtac ccgggggatcc cgttccaggg ctccacgtcc        40
325 <210> SEQ ID NO: 11
326 <211> LENGTH: 40
327 <212> TYPE: DNA
328 <213> ORGANISM: Artificial Sequence
330 <220> FEATURE:
331 <223> OTHER INFORMATION: primer 168Cfor
334 <400> SEQUENCE: 11
335 gcacatcccc ctttcgccag ccgatcacga gcttcagtcc          40
338 <210> SEQ ID NO: 12
339 <211> LENGTH: 20
340 <212> TYPE: DNA
341 <213> ORGANISM: Artificial Sequence
343 <220> FEATURE:
344 <223> OTHER INFORMATION: primer 168Crev
347 <400> SEQUENCE: 12
348 cgtcgtccgg gccagttcg                                20

```

VERIFICATION SUMMARY

DATE: 08/14/2006

PATENT APPLICATION: US/10/587,956

TIME: 13:18:36

Input Set : A:\Sequence Listing.TXT

Output Set: N:\CRF4\08142006\J587956.raw

L:8 M:270 C: Current Application Number differs, Replaced Current Application No

L:8 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:0 M:201 W: Mandatory field data missing, <130> FILE REFERENCE